

1. A recent study was conducted to explore how people’s dating preferences are affected by a prospective date's tendency to reveal (vs. hide) personal information. Each subject in the study was involved in an online dating service. They viewed two questionnaires that had supposedly been completed by two prospective dates: the *revealer* vs. the *hider*.

Revealer’s Questionnaire						Hider’s Questionnaire					
	Never	Once	Sometimes	Frequently	Choose not to answer		Never	Once	Sometimes	Frequently	Choose not to answer
Have you cheated on a tax return?				X		Have you cheated on a tax return?				X	
Have you made a false insurance claim?				X		Have you made a false insurance claim?				X	
Have you stolen anything worth more than \$100?				X		Have you stolen anything worth more than \$100?					X
Have you fantasized about hurting someone?				X		Have you fantasized about hurting someone?				X	
Have you hid an STD from a dating partner?				X		Have you hid an STD from a dating partner?					X

Note that the revealer answered all questions and admitted to negative behavior while doing so, while the hider selected “Choose not to answer” for two of the questions. The researchers hypothesized that dating prospects would be more likely to choose the revealer than the hider, in spite of the fact that the revealer admitted to more negative behavior (the theory is that potential dates are more afraid of those who hide information than they are of those who admit to negative information). **Of the 44 subjects tested, 28 preferred to date the revealer.**

The researchers’ hypothesis is as follows: *Potential dates more likely to choose the revealer than the hider.*

- a. If the researchers’ hypothesis is wrong and potential dates are no more likely to choose the revealer than they are to choose the hider, then how many of the 44 subjects do we expect to choose the revealer? (1 pt)

50% of 44 = 22

- b. Let π = the true proportion all potential dates that would choose the revealer. Calculate the statistic of interest in this study. (1 pt)

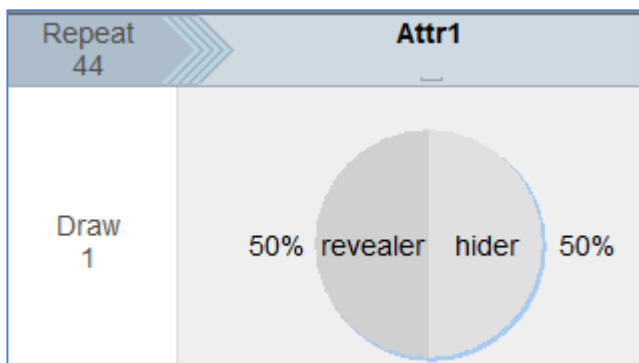
$\hat{\pi} = 28/44 = 63.6\%$

- c. Set up the null and alternative hypotheses to investigate the research hypothesis. (2 pts)

H_0 : Potential dates are no more likely to choose the revealer than the hider ($\pi = 0.50$)

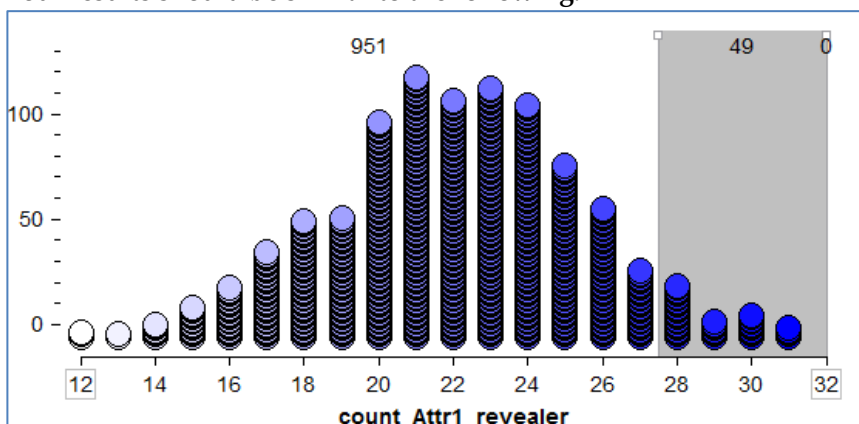
H_a : Potential dates are more likely to choose the revealer than the hider ($\pi > 0.50$)

- d. Next, set up a simulation in Tinkerplots that will allow you estimate a p-value for this hypothesis test. Sketch your spinner, including (1) the labels you used, (2) the probabilities attached to these labels, and (3) the Repeat value. (1 pt)



- e. Carry out 1,000 runs of your simulation study. Either include a detailed sketch or print a copy of your graph and attach it to your homework. (1 pt)

Your results should be similar to the following:



- f. Recall that in this study, 28 of the 44 subjects chose the revealer. Use the results of your simulation to estimate the p-value. (2 pts)

Estimated p-value: **Based the simulation study, estimate how often we see results such as 28 or more assuming the null hypothesis is true. From the results in part e, we estimate this to be $49/1000 = 0.049$.**

- g. Write a conclusion in the context of the problem to address whether or not the evidence supports the researchers' hypothesis (i.e., your conclusion should be about potential dates and their choice of the revealer vs. the hider). (2 pts)

Since the estimated p-value is below 0.05, the study provides evidence that potential dates are more likely to choose the revealer than the hider. Note that if your estimated p-value was between 0.05 and 0.10, you could either say "There is not enough evidence that potential dates are more likely to choose the revealer" or "There is marginal evidence that potential dates are more likely to choose the revealer."

2. Suppose a recent opinion poll surveyed 140 people who were randomly selected from a particular town and found that 90 of them disapproved of the way the mayor of that town was doing his job. A headline in the town's newspaper was subsequently written as follows: "Poll suggests that over 60% of the town's population disapproves of the mayor's job performance." Your job is to determine whether or not this headline is accurate.

a. Identify both the population of interest and the sample in this study. (1 pt)

Population: **All residents of this town**

Sample: **The 140 residents in this town who were actually surveyed**

b. Let π = the true proportion of this town's population who disapprove of their mayor's job performance. Calculate the statistic of interest in this study. (1 pt)

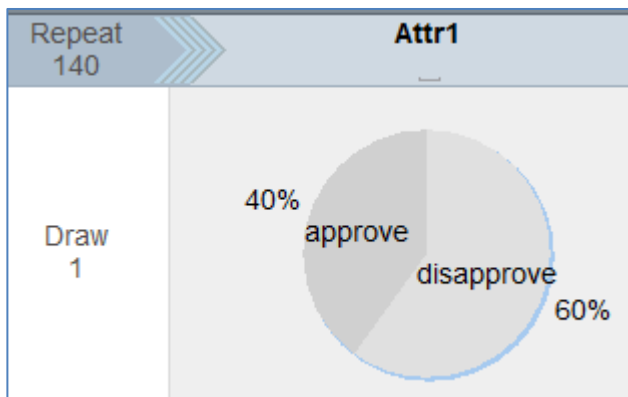
$$\hat{\pi} = 90/140 = 64.3\%$$

c. Set up the null and alternative hypotheses to test the research hypothesis. (2 pts)

H_0 : **60% of the town's residents disapprove of the mayor's job performance ($\pi = 0.60$)**

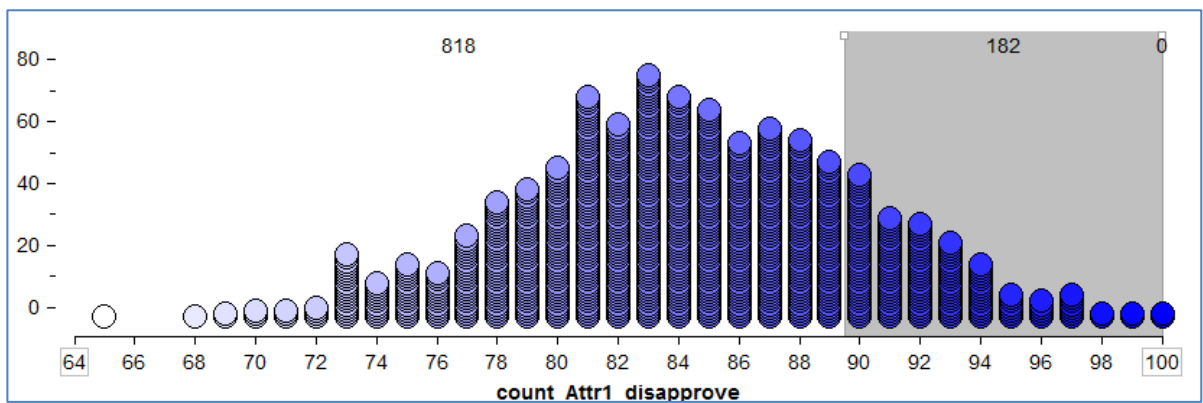
H_a : **Over 60% of the town's residents disapprove of the mayor's job performance ($\pi > 0.60$)**

d. Next, set up a simulation in Tinkerplots that will allow you estimate a p-value for investigating this research hypothesis. Sketch your spinner, including (1) the labels you used, (2) the probabilities attached to these labels, and (3) the Repeat value. (1 pt)



- e. Carry out 1,000 runs of your simulation study. Either include a detailed sketch or print a copy of your graph and attach it to your homework. (1 pt)

Your results should be similar to the following:



- f. Recall that in this poll, 90 of the 140 people surveyed said that they disapproved of the mayor's job performance. Use the results of your simulation to estimate the p-value. (2 pts)

Estimated p-value: **Based the simulation study, estimate how often we see results such as 90 or more assuming the null hypothesis is true. From the results in part e, we estimate this to be $182/1000 = 0.182$.**

- g. Write a conclusion in the context of the problem to address whether the headline is accurate or not. (2 pts)

Since the estimated p-value is above 0.05, this survey result does not provide enough evidence to conclude that over 60% of the town's residents disapprove of the mayor's job performance.

Even though more than 60% of those surveyed disapproved, our analysis indicates that a result such as $90/140 = 64.3\%$ disapproving was too likely to have happened just by chance, even if the true disapproval rate was not above 60%.